

10 / 529779

~~30 MAR 2005~~

PCT/DE 03 / 03246

BUNDESREPUBLIK DEUTSCHLAND

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)



REC'D 05 NOV 2003	
WIPO	PCT

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 45 490.6

Anmeldetag: 30. September 2002

Anmelder/Inhaber: Siemens Aktiengesellschaft, München/DE

Bezeichnung: Verfahren zur teilweisen Erhaltung der
Paketreihenfolge bei verbindungsloser
Paketvermittlung mit alternativem Routing

IPC: H 04 L 12/56

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der
ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 16. Oktober 2003
Deutsches Patent- und Markenamt
Der Präsident

Im Auftrag

Faust

A 9161
03/00
EDV-L

BEST AVAILABLE COPY



Beschreibung

Verfahren zur teilweisen Erhaltung der Paketreihenfolge bei verbindungsloser Paketvermittlung mit alternativem Routing

5

Der Anmeldungsgegenstand betrifft ein Verfahren zur Erhaltung der Paketreihenfolge bei verbindungsloser Paketvermittlung mit alternativem Routing für ein Netz mit mehreren Routern.

- 10 Netze mit verbindungsloser Paketvermittlung (z.B. das heutige Internet) haben im Regelfall keine Möglichkeit, beim Transport durch das Netz die Reihenfolge von Paketen zu erhalten, d.h. am Ausgang des Netzes dieselbe Reihenfolge anzubieten wie am Eingang, wenn innerhalb des Netzes beispielsweise zur
- 15 Lastverteilung die Route zu einem Ziel für jedes Paket individuell gewählt werden kann.

- Fehler in der Paketreihenfolge können beispielsweise den Durchsatz von Verbindungen unnötigerweise verringern, wenn
- 20 dieser durch das Protokoll TCP (Transmission Control Protocol) geregelt wird.

Um das Netz nicht zu überlasten und eine faire Aufteilung der gesamten Bitrate auf viele Verbindungen zu erreichen, regelt ein TCP-Sender seine Senderate (durch Verringerung des Sendefensters) nach Erkennen eines Paketverlustes herunter. Auch eine im Netz vertauschte Paketreihenfolge führt in der Praxis zu wiederholten Bestätigungen mit derselben Reihenfolgenummer, so dass TCP auch hier die Rate reduziert.

30

- Alternatives Routing auf Paketebene, d. h. Verkehrsverteilung von Paketen eines Flows, wird aus den o.g. Gründen heute in der Regel nicht eingesetzt. Um die o.g. Probleme zu vermeiden, wird beispielsweise bei MPLS (Multi Protocol Label Switching) alternatives Routing auf der Ebene (aggregierter)
- 35 Flows vorgeschlagen, d.h. alle Pakete, die zur selben Verbindung gehören oder zwischen demselben Paar von Netzknoten aus-

- getauscht werden, werden auf demselben Weg durch das Netz geschickt. Hierzu muss allerdings in jedem Netzknoten entsprechende Verbindungsinformation hinterlegt werden, z. B. indem Pfade konfiguriert werden (statisch), oder indem für jede
- 5 Verbindung zunächst ein Pfad aufgebaut wird (dynamisch, aber mit viel Aufwand verbunden und daher nicht unbedingt skalierbar für große Netze). Die Anzahl der abzuspeichernden Flows hängt hier sehr stark von der Dauer der Flows ab und kann bei langen Flows mit jeweils wenig Verkehr sehr groß werden.
- 10 Desweiteren können am Netzausgang Einrichtungen eingesetzt werden, die die Paketreihenfolge wieder herstellen, was allerdings in IP (Internet Protocol)-Netzen eine nicht-triviale Aufgabe ist, da IP-Pakete in der Regel keine für solche Zwecke verwendbare Reihenfolgennummer besitzen. Das „Identity“-
- 15 Feld im Paketkopf identifiziert ein Paket zwar eindeutig, wird aber nicht unbedingt innerhalb jeder TCP-Verbindung oder jeder UDP (User datagram protocol) -Assoziation jeweils um 1 erhöht. Um die auf Oktett-Ebene angegebene TCP-„Sequence Nummer“ auszuwerten, muss der Paketkopf weiter ausgewertet werden, da sich diese Nummer von einem TCP-Segment zum nächsten um die Anzahl der Bytes im Segment erhöht. Da die Segmente
- 20 außerdem auch innerhalb einer Verbindung unterschiedliche Mengen von Nutzinformation tragen können, kann eine Resequenzierereinrichtung nicht wissen, wieviele Pakete zwischen zwei anderen empfangenen Paketen noch fehlen, wenn deren Sequenznummern nicht aneinander anschließen. Zusätzlich würden beim Einsatz einer Resequenzierereinrichtung Paketverluste eine Verzögerung des Ausspielens der Pakete bewirken und damit den
- 25 „Fast Retransmit“-Mechanismus von TCP außer Kraft setzen, was die Bandbreitenregelung in TCP zum starken Herabregeln veranlassen würde und damit im Vergleich zum Abliefern außerhalb der Reihenfolge keinen Vorteil brächte.
- 30
- 35 Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren anzugeben, das in einem Netz mit multiplen Wegemöglichkeiten

durch Paketüberholungen bedingte Performance-Degradationen vermindert.

Die Aufgabe wird durch die Merkmale des Anspruchs 1 oder des
5 Anspruchs 4 gelöst.

Durch die Erfindung wird die Häufigkeit von Paketüberholungen, insbesondere für hochbitratige Verbindungen, stark vermindert. Die Häufigkeit von Vertauschungen in der Paketreihenfolge wird durch die o.g. technischen Merkmale verringert. Die maximale Anzahl von Flow-Einträgen in der Flowtable FT ist durch die Anzahl von im Router abzuspeichernden Paketen vorgegeben. Durch die Beschränkung auf im Router gespeicherte Pakete wird somit die Menge an Zustandsinformationen im Router gegenüber Lösungen wie MPLS oder IP Switching, die für jeden existierenden Flow einen Zustand halten müssen, stark verringert. Außerdem wird im Gegensatz zu MPLS oder IP Switching keine Signalisierung zwischen den Netzknoten benötigt, so dass insbesondere bei kurzen Flows keine unnötige Verzögerung auftritt. Die Beschränkung auf die kurze Lebensdauer der Zustandsinformation hat darüberhinaus den Vorteil, dass die Flexibilität beim alternativen Routing zur Lastverteilung im Netz gewährleistet bleibt, so dass ein Kompromiss zwischen der absoluten Einhaltung der Paketreihenfolge und optimaler Lastverteilung erreicht werden kann. Verbindungen, die mit einer hohen Rate senden und von denen stets mindestens ein Paket im Router zwischengespeichert ist, werden keine Reihenfolgevertauschungen erfahren. Verbindungen, in denen nur selten ein Paket gesendet wird, werden auch keine Probleme bekommen, wenn die Laufzeitunterschiede zwischen den verschiedenen im Netz gewählten Pfaden im Vergleich zur Zeit zwischen zwei Paketen klein sind. Insbesondere für Verbindungen, in denen Daten büschelweise gesendet werden (z.B. World Wide Web) ist die beschriebene Lösung daher vorteilhaft.

35

Vorteilhafte Weiterbildungen der Erfindung sind in den Unteransprüchen angegeben.

Die Erfindung wird im folgenden als Ausführungsbeispiel in einem zum Verständnis erforderlichen Umfang anhand von Figuren näher erläutert. Dabei zeigen:

- 5 Fig 1 eine vereinfachte Darstellung eines IP-Netzes,
 Fig 2 eine schematische Darstellung eines IP-Routers,
 Fig 3 eine schematische Darstellung eines erfindungsgemäßen
 IP-Routers und
10 Fig 4 eine schematische Darstellung des Inhaltes der Flow-
 Tabelle FT.

In den Figuren bezeichnen gleiche Bezeichnungen gleiche Elemente.

- 15 In Figur 1 ist ein einfaches Netz dargestellt, in dem zwei
 Endeinrichtungen E1 und E2 über mehrere Wege miteinander ver-
 bunden sind, wobei die Router R1 bis R5 eine verbindungslose
 Paketvermittlung zwischen den Links (Leitungen) L0, L1, L2,
20 ... L7 vornehmen sollen. Figur 2 zeigt einen Teil des IP-
 Routers R1, wie er nach dem Stand der Technik aufgebaut ist,
 für eine Übermittlungsrichtung (von L0 hin zu L1 und L2).
 Wenn ein Paket ankommt, wird es klassifiziert, die Ziel-IP-
 Adresse wird ausgelesen und für diese Adresse wird aus der
25 Routing-Tabelle RT der nächste Router auf dem Weg zum Ziel
 bestimmt. Die Routing-Tabelle erhält aktuelle Routing-
 Informationen vom Routing-Protokoll-Prozessor RP, der über
 ein Routing-Protokoll Erreichbarkeitsinformationen mit ande-
 ren Routern austauscht. In der Regel wird der „shortest
30 path“, also der (nach einer vorgebbaren Metrik) kürzeste Weg
 zum Ziel als einziger Weg in die Routing-Tabelle RT eingetra-
 gen.

- Im Falle der Lastverteilung auf mehrere alternative Routen
35 ist die Routing-Tabelle erweitert und enthält neben dem
 nächsten Knoten auf dem kürzesten Pfad noch weitere nächste
 Knoten für weitere zulässige Pfade zum Ziel. Bei jedem ankam-

menden Paket kann nun aufgrund eines Lastverteilungsalgorithmus ein zulässiger Ausgangspfad zum Ziel gewählt werden, an den das Paket dann weitergeleitet wird.

- 5 Erfindungsgemäß wird vorgeschlagen, im Router eine Tabelle von Flow- oder Verbindungsinformationen zu halten (im Folgenden nur noch als Flow-Information FI bezeichnet), die für jedes im Router befindliche (also in einer der Warteschlangen Q1, Q2 oder Q3 zwischengespeicherte oder gerade im Koppelnetz
- 10 vermittelte) Paket die gewählte Route speichert. Verlässt (bei dieser Ausführungsform) das Paket den Router, dann wird die Information wieder gelöscht. Kommt ein neues Paket mit derselben FI in den Router, dann wird es auf demselben Pfad weitergeleitet wie das letzte Paket mit derselben FI.

- 15 Die Entscheidung, auf welchem der alternativen Wege ein Paket weitergeleitet wird, wird demnach nur dann neu getroffen, wenn kein Paket mit derselben FI wie ein neu angekommenes Paket sich bereits im Router befindet. Dadurch wird die Häufig-
- 20 keit von Paketüberholungen für hochbitratige Verbindungen stark vermindert.

- Der entsprechende Router in Figur 3 enthält zusätzlich zu den Komponenten des Routers aus Figur 2 eine Flow-Tabelle FT, in der für alle noch im Router befindlichen und bereits klassifizierten Pakete der ausgewählte next hop abgelegt ist. Für jedes neu eintreffende Paket wird die Zugehörigkeit zu einem der Flows in der FT geprüft. Falls ein Paket des entsprechenden Flows in der FT verzeichnet ist, wird dieselbe Auswahl
- 30 des next hop auch für das neue Paket getroffen. Falls kein Paket desselben Flows in der FT verzeichnet ist, wird anhand der Regeln des alternativen Routing und der Lastverteilung für dieses Paket ein next hop ausgewählt, das Paket in Richtung dieses nächsten Knotens weitergeleitet und die Flow-
- 35 Information zusammen mit dem gewählten next hop in der FT abgespeichert. Figur 4 zeigt als Beispiel, wie eine solche FT aussehen kann. Die FT enthält für jeden Flow i, von dem sich

Pakete in einer Ausgangswarteschlange des Routers befinden, die Anzahl n_i der Pakete in den Warteschlangen, die Flowkennungsinformationen (source IP, destination IP, source Port, dest. Port, Protocol) und den für diesen Flow gewählten next Hop. Der Paketzähler für jeden Flow wird mit jedem ankommenden Paket für diesen Flow um 1 erhöht und mit jedem aus dem Router abgehenden Paket für diesen Flow um 1 verringert. Wenn der Zähler dabei den Wert 0 erreicht, wird der Eintrag aus der Tabelle gelöscht.

10

Weitere Ausgestaltungen:

1. Das Prinzip kann für jede Warteschlange einzeln, auf eine Teilmenge oder auf alle Puffer in einem Gerät angewendet werden, falls ein IP-Router beispielsweise Ein- und Ausgangspuffer oder eine Kombination solcher Puffer mit einem zentralen Puffer einsetzt. Es sind folgende Alternativen möglich:

a) getrennte FT und getrennte Paketzählung. In diesem Fall bezieht sich die FT nur auf die Warteschlange, an deren Ausgang die Entscheidung zur Weiterleitung auf einen bestimmten Weg getroffen wird. Eventuell dahinter angeordnete Ausgangspuffer und die Pakete darin haben keinen Einfluss mehr auf die Wegeentscheidung für neue Pakete.

b) gemeinsame FT und getrennte Paketzählung. Die FT enthält in diesem Fall einen Paketzähler pro Warteschlange, der jeweils bei Ein- und Austritt eines Paketes in die / aus der Warteschlange aktualisiert wird. Die Forwarding-Entscheidung wird per Flow gespeichert.

c) gemeinsame FT und gemeinsame Paketzählung. Die FT ist nach Figur 4 strukturiert, wobei n_i sich auf die Summe aller Pakete des Flows i in allen betrachteten Warteschlangen bezieht.

Die Forwarding-Entscheidung wird dabei auch durch die Pakete eines Flows noch beeinflusst, die die Entscheidungsstelle bereits passiert haben. Diese Option b)/c) ist gegenüber a) vorzuziehen.

2. IP-Router haben in der Regel eine Ausgangswarteschlange pro Ausgangslink, wobei der Ausgangslink eine physikalische Netzverbindung oder ein logischer Kanal innerhalb einer physikalischen Verbindung sein kann (z.B. ein ATM-VP (Asynchronous Transfer Mode-Virtual Path) oder -VC (Virtual Channel), eine Wellenlänge oder ein STM (Synchronous Transport Modul) -Kanal). In Backbone-Routern ist in der Regel jedem dieser Kanäle genau ein nächster IP-Router zugeordnet. In lokalen Netzen kann dagegen ein Ausgangskanal auch mehrere nächste IP-Router erreichen, wenn der Kanal z.B. ein shared medium (Ethernet ö.ä.) ist. In diesem Fall gibt es die Optionen, entweder den Ausgangskanal oder - wie in Figur 4 angedeutet - den „next hop“ für einen Flow in der Flow-Tabelle FT einzutragen. Letztere Möglichkeit erscheint sinnvoller; aus anderen Gründen (z.B. interne Struktur des Routers) kann es aber nötig sein, den Ausgangskanal als Ersatz für den „next hop“ einzutragen.
3. In der Beschreibung zu Figur 3 ist vorgesehen, Flows aus der Flow-Tabelle auszutragen, sobald sich kein entsprechendes Paket mehr im Router befindet. Alternativ kann auch ein Ausaltern vorgesehen werden, bei dem in der Flow-Tabelle FT in Figur 4 anstelle der Anzahl n_i von Paketen eines Flows ein Zeitstempel für die letzte Paketankunft gespeichert wird. Die Einträge werden dann periodisch oder nach Ablauf einer Zeitschranke nach der Ankunft aus der Tabelle ausgetragen, wenn der Zeitpunkt, an dem das letzte Paket eines Flows beobachtet wurde, bereits mindestens eine vorzugebende Zeit zurückliegt.
4. Falls ein Router mit mehreren Verkehrsklassen umgeht, kann das Verfahren für alle oder nur für einen Teil der Verkehrsklassen eingesetzt werden.
5. Eine Zeitschranke nach Option 3 kann adaptiv in Abhängigkeit anderer Parameter eingestellt werden. Hierfür kommen insbesondere Parameter in Frage, die die Verkehrsverteilung bestimmen (z.B. die Häufigkeit für die Wahl eines Alternativweges).

Patentansprüche

1. Verfahren zur Erhaltung der Paketreihenfolge bei verbindungsloser Paketvermittlung mit alternativem Routing für ein
5 Netz mit mehreren Routern
demzufolge
 - für ein in einem Router ankommendes Paket eine Route ausgewählt wird,
 - für den zugehörigen flow des Pakets die ausgewählte Route
10 in dem Router gespeichert wird,
 - das Paket auf die ausgewählte Route weitergeleitet wird,
 - ein neu ankommendes Paket, das dem selben flow zugehört, auf die gespeicherte Route weitergeleitet wird und
 - für den Fall, dass sich für einen flow kein zugehöriges
15 Paket in dem Router befindet, die ausgewählte Route in dem Router gelöscht wird.
2. Verfahren nach Anspruch 1
dadurch gekennzeichnet, dass
20 als Kriterium, ob sich für einen flow ein Paket in dem Router befindet, die Anzahl der dem flow zugehörigen Pakete in einer Warteschlange des Routers ausgewertet wird.
3. Verfahren nach Anspruch 2
25 dadurch gekennzeichnet, dass
als Kriterium, ob sich für einen flow ein Paket in dem Router befindet, die Anzahl der dem flow zugehörigen Pakete in genau einer Warteschlange des Routers ausgewertet wird.
- 30 4. Verfahren zur Erhaltung der Paketreihenfolge bei verbindungsloser Paketvermittlung mit alternativem Routing für ein Netz mit mehreren Routern
demzufolge
 - für ein in einem Router ankommendes Paket der Ankunfts-
35 zeitpunkt in dem Router gespeichert wird,
 - für das in dem Router ankommende Paket eine Route ausgewählt wird,

- für den zugehörigen flow des Pakets die ausgewählte Route in dem Router gespeichert wird,
- das Paket auf die ausgewählte Route weitergeleitet wird,
- für den zugehörigen flow eines neu ankommenden Pakets der Ankunftszeitpunkt aktualisiert wird
- 5 - ein neu ankommendes Paket, für dessen flow eine Route abgespeichert ist, auf die gespeicherte Route weitergeleitet wird und
- 10 - für den Fall, dass seit dem Ankunftszeitpunkt des zuletzt angekommenen Pakets eines flows eine vorgegebene Zeitdauer abgelaufen ist, die Abspeicherungen für den flow gelöscht werden.

5. Verfahren nach Anspruch 4
15 dadurch gekennzeichnet, dass
die Abspeicherungen für die einzelnen flows periodisch gelöscht werden.

6. Verfahren nach Anspruch 4 oder 5
20 dadurch gekennzeichnet, dass
die vorgegebene Zeitdauer nach Massgabe eines die Verkehrsverteilung bestimmenden Kriteriums eingestellt wird.

7. Verfahren nach einem der vorstehenden Ansprüche
dadurch gekennzeichnet, dass
die Abspeicherungen für die einzelnen flows in einer flowtable (FT) erfolgen.

Zusammenfassung

Verfahren zur teilweisen Erhaltung der Paketreihenfolge bei verbindungsloser Paketvermittlung mit alternativem Routing

5

In einem verbindungslosen Netz mit mehreren Routern, bei dem die Pakete eines Flows über mehrere Wegemöglichkeiten weiterleitbar sind, wird zur Erhöhung der Einhaltung der Paketreihenfolge vorgeschlagen ein in einem Router ankommendes Paket auf dieselbe Route weiterzuleiten wie ein gerade sich im Router befindliches, zum selben Flow gehörendes Paket oder ein in dem Router ankommendes Paket, das innerhalb eines Zeitfensters in dem Router ankommt, das von einem zum selben Flow gehörenden Paket geöffnet wurde, auf derselben Route weitergeleitet wird wie das das Fenster öffnende Paket.

10

15

Fig 3

Bezugszeichenliste

	L0, L1, L2	Links, Verbindungsleitungen (nur eine Richtung gezeigt)
5	KN	Koppelnetz
	Q1, Q2, Q3	Warteschlangen
	RT	Routing Table
	RP	Routing-Protokoll-Prozessor
	FT	Flow Table
10	UDP	User datagram protocol

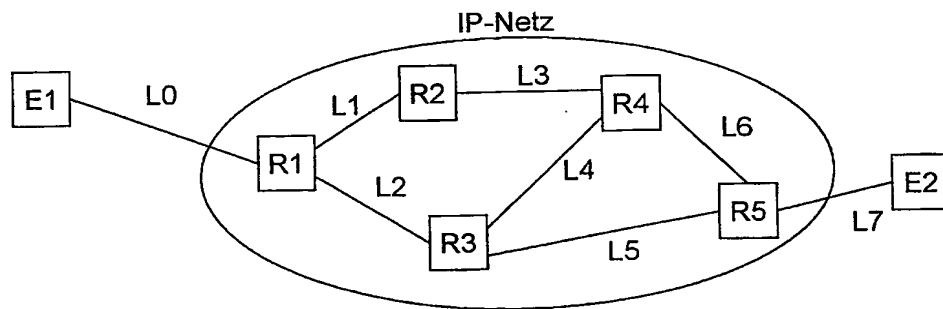


FIG 1

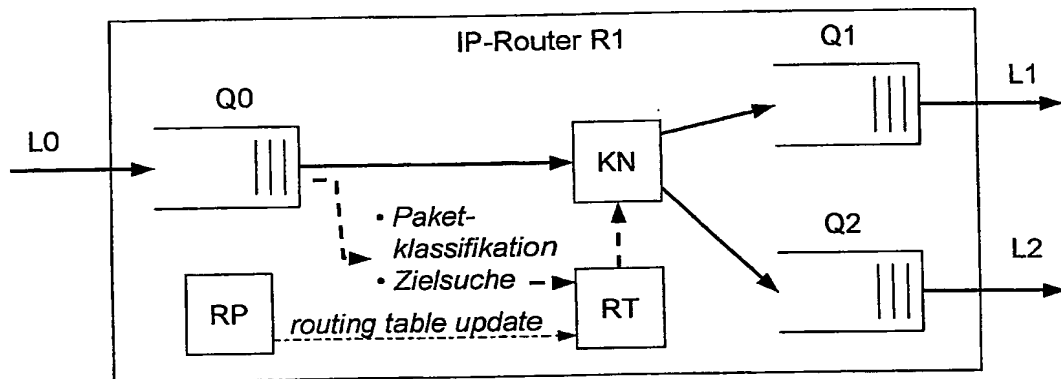


FIG 2

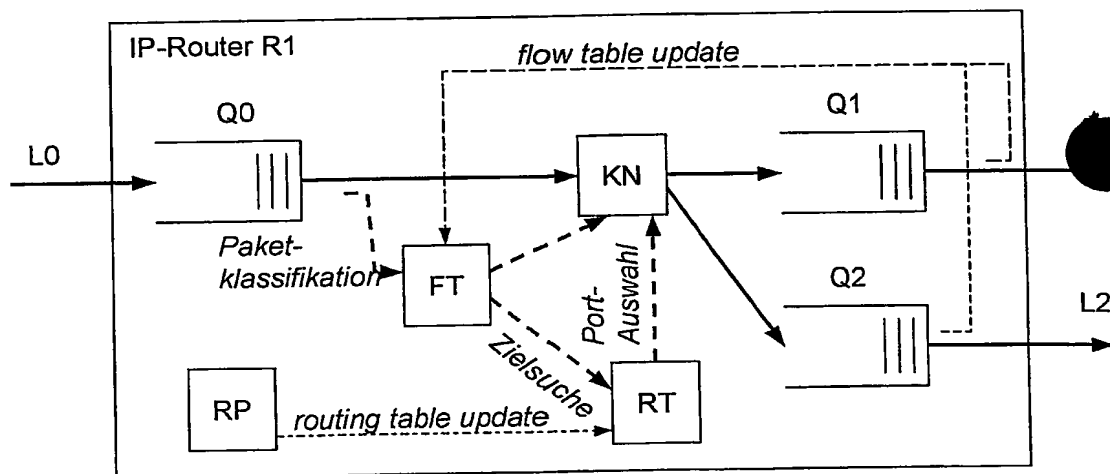


FIG 3

IP1:Port1	IP2:Port2	Protocol	n_i	nextHop
120.23.14.15:80	144.6.32.23:1051	TCP	3	R2
110.3.1.52:1200	144.6.32.23:7552	TCP	2	R2
151.2.41.125:23	144.6.32.23:2353	TCP	1	R3
120.23.14.15:80	144.6.32.23:1052	TCP	10	R2

FIG 4

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.